

Analysis of circulant embedding methods for sampling stationary random fields*

I.G. Graham[†] F.Y. Kuo[‡] D. Nuyens[§] R. Scheichl[¶]
I.H. Sloan^{||}

March 20, 2018

Abstract

A standard problem in uncertainty quantification and in computational statistics is the sampling of stationary Gaussian random fields with given covariance in a d -dimensional (physical) domain. In many applications it is sufficient to perform the sampling on a regular grid on a cube enclosing the physical domain, in which case the corresponding covariance matrix is nested block Toeplitz. After extension to a nested block circulant matrix, this can be diagonalised by FFT – the “circulant embedding method”. Provided the circulant matrix is positive definite, this provides a finite expansion of the field in terms of a deterministic basis, with coefficients given by i.i.d. standard normals. In this paper we prove, under mild conditions, that the positive definiteness of the circulant matrix is always guaranteed, provided the enclosing cube is sufficiently large. We examine in detail the case of the Matérn covariance, and prove (for fixed correlation length) that, as $h_0 \rightarrow 0$, positive definiteness is guaranteed when the random field is sampled on a cube of size order $(1 + \nu^{1/2} \log h_0^{-1})$ times larger than the size of the physical domain. (Here h_0 is the mesh spacing of the regular grid and ν the Matérn smoothness parameter.) We show that the sampling cube can become smaller as the correlation length decreases when h_0 and ν are fixed. Our results

*March 20, 2018. The authors acknowledge financial support from the Australian Research Council (FT130100655; DP150101770), the KU Leuven research fund (OT:3E130287; C3:3E150478), the Taiwanese National Center for Theoretical Sciences’ Mathematics Division, and the Statistical and Applied Mathematical Sciences Institute (SAMSI) under its 2017 Program on Quasi-Monte Carlo and High-Dimensional Sampling Methods for Applied Mathematics.

[†]Dept. of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK (I.G.Graham@bath.ac.uk).

[‡]School of Mathematics and Statistics, University of NSW, Sydney NSW 2052, Australia (f.kuo@unsw.edu.au).

[§]Dept. of Computer Science, KU Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium (dirk.nuyens@cs.kuleuven.be).

[¶]Dept. of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK (R.Scheichl@bath.ac.uk).

^{||}School of Mathematics and Statistics, University of NSW, Sydney NSW 2052, Australia (i.sloan@unsw.edu.au).

are confirmed by numerical experiments. We prove several results about the decay of the eigenvalues of the circulant matrix. These lead to the conjecture, verified by numerical experiment, that they decay with the same rate as the Karhunen–Loève eigenvalues of the covariance operator. The method analysed here complements the numerical experiments for uncertainty quantification in porous media problems in an earlier paper by the same authors in *J. Comp. Physics*. 230 (2011), pp. 3668–3694.

Keywords: Gaussian Random Fields, Circulant Embedding, Statistical Homogeneity, Matérn Covariance, Fast Fourier Transform, Analysis

AMS: 60G10, 60G60, 65C05, 65C60

1 Introduction

In recent years there has been a huge growth in interest in uncertainty quantification (UQ) for physical models involving partial differential equations. In this context, the forward problem of UQ consists of describing the statistics of outputs (solutions) of a PDE model, given statistical assumptions on its inputs (e.g., its coefficients). This has led to the widespread study of model PDE problems where the coefficients are given as random fields. Various flavours of solution method (e.g., Stochastic Galerkin or various sampling methods) have been formulated under the assumption that the random coefficient field has a separable expansion in physical/probability space – for example a suitably truncated Karhunen–Loève (KL) expansion, [10], [18].

In its standard form, the KL expansion is in principle infinite, and requires the computation of the (infinitely many) eigenpairs of the integral operator with kernel given by the covariance of the field. It has to be truncated to be computable, even when the field is only required at a finite set of physical spatial points. However if it is known from the outset that the coefficient field is only required at a finite set of physical points (as is the case in typical implementations of finite element methods for solving the PDE), then a different point of view emerges. The required field is now a random vector; in the Gaussian case it is characterised by its mean and covariance matrix. A real, for example Cholesky, factorization of the covariance matrix provides a finite separable expansion of the random vector, with no need for any truncation.

In the paper [11], the authors proposed a practical algorithm for solving a class of elliptic PDEs with coefficients given by statistically homogeneous lognormal random fields with low regularity. Lognormal random fields are commonly used in applications, for example in hydrology (see, e.g., [19, 20] and the references there).

The PDE was solved by piecewise linear finite elements on a uniform grid. The stiffness matrix was obtained by an appropriate quadrature rule, with the field values at the quadrature points being obtained via a factorization of the covariance matrix using the circulant embedding technique described below. The method was found to be effective even for problems with high stochastic dimension, but [11] did not contain a convergence analysis of the algorithm. The main purpose of the present

paper is to provide an analysis for the circulant embedding part of the algorithm of [11]. The analysis of the corresponding uncertainty quantification algorithm for the PDE is done in [12].

We consider here the fast evaluation of a Gaussian random field $Z(\mathbf{x}, \omega)$ with prescribed mean $\bar{Z}(\mathbf{x})$ and covariance

$$r_{\text{cov}}(\mathbf{x}, \mathbf{x}') := \mathbb{E}[(Z(\mathbf{x}, \cdot) - \bar{Z}(\mathbf{x}))(Z(\mathbf{x}', \cdot) - \bar{Z}(\mathbf{x}'))], \quad (1.1)$$

where the expectation is with respect to the Gaussian measure. Throughout we will assume that Z is *stationary* (see, e.g., [2, p. 24]), i.e., its covariance function satisfies

$$r_{\text{cov}}(\mathbf{x}, \mathbf{x}') = \rho(\mathbf{x} - \mathbf{x}'), \quad (1.2)$$

for some function $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$. Note that we assume here that ρ is defined on all of \mathbb{R}^d , as it is in many applications, although strictly speaking we only need ρ to be defined on a sufficiently large ball. Note also that (1.1) and (1.2) imply that ρ is symmetric, i.e.,

$$\rho(\mathbf{x}) = \rho(-\mathbf{x}), \quad \text{for all } \mathbf{x} \in \mathbb{R}^d, \quad (1.3)$$

and that ρ is also positive semidefinite (see §2.2). Further assumptions on ρ will be given below. A particular case, to be discussed extensively, is the Matérn covariance defined in Example 2.7 below.

We shall consider the problem of evaluating $Z(\mathbf{x}, \omega)$ at a uniform grid of

$$M = (m_0 + 1)^d$$

points on the d -dimensional unit cube $[0, 1]^d$, with integer m_0 fixed, and with grid spacing $h_0 := 1/m_0$. (The extension to general tensor product grids is straightforward and not discussed here.) Denoting the grid points by $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$, we wish to obtain samples of the random vector:

$$\mathbf{Z}(\omega) := (Z(\mathbf{x}_1, \omega), \dots, Z(\mathbf{x}_M, \omega))^\top.$$

This is a Gaussian random vector with mean $\bar{\mathbf{Z}} := (\bar{Z}(\mathbf{x}_1), \dots, \bar{Z}(\mathbf{x}_M))^\top$ and a positive semidefinite covariance matrix

$$R = [\rho(\mathbf{x}_i - \mathbf{x}_j)]_{i,j=1}^M. \quad (1.4)$$

Because of its finite length, $\mathbf{Z}(\omega)$ can be expressed exactly (but not uniquely) as a linear combination of a finite number of i.i.d. standard normals, i.e., as

$$\mathbf{Z}(\omega) = B\mathbf{Y}(\omega) + \bar{\mathbf{Z}}, \quad \text{where } \mathbf{Y} \sim \mathcal{N}(\mathbf{0}, I_{s \times s}). \quad (1.5)$$

for some real $M \times s$ matrix B with $s \geq M$ satisfying

$$R = BB^\top. \quad (1.6)$$

To verify this construction, simply note that (1.5) and (1.6) imply

$$\mathbb{E}[(\mathbf{Z} - \bar{\mathbf{Z}})(\mathbf{Z} - \bar{\mathbf{Z}})^\top] = \mathbb{E}[B\mathbf{Y}\mathbf{Y}^\top B^\top] = B\mathbb{E}[\mathbf{Y}\mathbf{Y}^\top]B^\top = BB^\top = R,$$

so ensuring that (1.1) is satisfied on the discrete grid.

In principle the factorization (1.6) could be computed via a Cholesky factorization (or even a spectral decomposition) of R with $s = M$, but this is likely to be prohibitively expensive, since R is large and dense. However, under appropriate ordering of the indices, R is a nested block Toeplitz matrix and, as we will explain below, R can be embedded in a bigger $s \times s$ nested block circulant matrix whose spectral decomposition can be rapidly computed using FFT with $\mathcal{O}(s \log s)$ complexity. A subtle but vital point is that while the covariance matrix R is automatically positive semidefinite, the extension to a larger nested block circulant matrix may lose definiteness, yet the nested block circulant matrix must be at least positive semidefinite for the algorithm to work. Small deviations from positive semidefiniteness are acceptable if one is prepared to accept the incurred errors from omitted negative eigenvalues. That error can be controlled via an a posteriori bound in terms of the negative eigenvalues (see, e.g., [18, §6.5]). In the present paper, however, we insist on positive definiteness.

The principle of our extension of the covariance matrix R is as follows. (Full details are in §2.1 below.) We first embed the unit cube $[0, 1]^d$ in a larger cube $[0, \ell]^d$ with side length $\ell = mh_0 \geq 1$ for some integer $m \geq m_0 = 1/h_0$. Note that ρ is automatically defined on $[0, \ell]^d$, since it is defined on all of \mathbb{R}^d . Then we construct a 2ℓ -periodic even symmetric extension of ρ on $[0, 2\ell]^d$, called ρ^{ext} , see (2.4) below, that coincides with ρ on $[0, \ell]^d$. The extended $s \times s$ matrix with

$$s = (2m)^d \tag{1.7}$$

is then obtained by the analogue of formula (1.4), with ρ replaced by ρ^{ext} . In Theorem 2.3, we show, under quite general conditions, that if ℓ (equivalently m) is chosen large enough, this extension is necessarily positive definite. The algorithm used in practice (Algorithm 1 in §2 below) extends ℓ cautiously through a sequence of increments in m until positive definiteness is achieved.

To know that the resulting algorithm is efficient, we need a lower bound on the value of ℓ needed to achieve positive definiteness. Our second set of theoretical results provides such bounds for the important Matérn class of covariance functions, defined in (2.21) below. In Theorem 2.9 we show that positive definiteness is always achieved with

$$\ell/\lambda \geq C_1 + C_2 \nu^{1/2} \log(\max\{\lambda/h_0, \nu^{1/2}\}) , \tag{1.8}$$

where λ is a parameter with the unit of length (the “correlation length”), $\nu < \infty$ is the Matérn smoothness parameter and C_1, C_2 are constants independent of h_0, ℓ, λ, ν and variance σ^2 . Thus the required ℓ grows very slowly in $1/h_0$ and can get smaller as the correlation length decreases. There is some growth as ν increases. However, the less smooth fields with ν small, and with small correlation length λ , are the ones often found in applications – see the references in [11]. In Theorem 2.11 we discuss the same question in the Gaussian case ($\nu = \infty$). This theorem shows, for example, that if λ and h_0 both decrease but λ/h_0 is kept fixed (i.e., a fixed number of grid points per unit correlation length), then the minimum value of ℓ needed for positive definiteness can decrease linearly in λ .

An additional benefit of the spectral decomposition obtained by applying the FFT to the circulant matrix is that it allows us to determine empirically which variables are the most important in the system. This gives an ordering of the variables, which can be used to drive the design of the Quasi-Monte Carlo algorithms (see, e.g., [12]).

While circulant embedding techniques are well-known in the computational statistics literature (e.g., [5, 6, 7, 17, 18]), there is relatively little theoretical analysis of this technique, the best existing references being Chan and Wood [5] and Dietrich and Newsam [7]. First, [5] provides a theorem in general dimension d identifying conditions on the covariance function which ensure that the circulant extension (which we shall describe below) is positive definite for some sufficiently large ℓ . The condition is similar to that provided in our Theorem 2.3 below, except that an additional assumption on the discrete Fourier transform (the “spectral density”) of ρ is required in [5]. In our work we require positivity of the *continuous* Fourier transform which is automatically satisfied via Bochner’s theorem, due to the fact that ρ is a covariance kernel. On the other hand Dietrich and Newsam [7] provide more detailed information about the behaviour of the algorithm by restricting the theory to a 1-dimensional domain. They show that the discrete Fourier transform of ρ will be positive when r_{cov} is convex, decreasing and non-negative. This automatically proves the success of the algorithm for certain covariance kernels. However, two covariance kernels that are not covered by this theory are the “Whittle” covariance and the Gaussian covariance. These belong to the Matérn class introduced in Example 2.7 below, with Matérn parameter $\nu = 1$ and $\nu = \infty$ respectively. Our theory covers the whole Matérn family and also describes the behaviour of the embedding algorithm with respect to the parameters ν (Matérn parameter) and λ (correlation length) as well as the mesh size h_0 . Finally we note that [7] also describes embedding strategies which are more general than those which we describe and analyse here, but without theory.

We may compare the present approach with other methods for the generation of Gaussian random fields. The principal limitations of the present approach are that it requires the covariance function to be stationary, and generates the random field only on a uniform tensor product grid. The requirement of uniformity of the grid is not very restrictive for PDE applications, as we show in [12] where we take the interpolation error into account for the error analysis. In [9], the uniform grid requirement is removed but at the price of approximating the covariance matrix by an H^2 -matrix, and then obtaining an approximate square root of that H^2 -matrix. The use of H -matrices to promote the efficient computation of KL eigenfunctions has previously been explored in [8, 16, 14]. In [21], fast multipole methods were promoted to achieve the same objective. In all these approaches, there is an inevitable need to truncate the KL series, but on the other hand there is no essential restriction to stationary covariance functions. We mention also the recent paper [3] in which a related periodization procedure is applied (this time using a smooth cut-off function), facilitating wavelet expansions of Z .

An interesting paper closely related to the present paper is [14], where the authors propose and analyse an approximate, pivoted Cholesky factorisation followed by a small eigendecomposition, which provides in many cases a very efficient and accurate

way to approximate the random field. This approach does not require stationarity or a uniform grid, but for efficiency reasons the factorisation needs to be truncated after $K \ll M$ steps, where in the simplest case M is the number of grid points as in the present paper. For fixed K , the accuracy of the truncated expansion depends on the decay of the KL eigenvalues and the cost is $\mathcal{O}(K^2 M)$. When the KL eigenvalues do not decay sufficiently fast, this limits the possible accuracy or leads to prohibitively high costs.

The structure of this paper is as follows. The circulant embedding algorithm is described in §2.1. The general theory of positive definiteness of the circulant extension of R is given in §2.2. General isotropic covariances (of which the Matérn is an example) are studied in §2.3, with the estimate (1.8) proved in Theorem 2.9. A key ingredient in the theory of the related paper [12] is an estimate on the rate of decay of the eigenvalues of the the circulant extension of the original covariance matrix. We discuss this question in §3 where we prove several results about the eigenvalues of the circulant matrix, leading to the conjecture that they decay at the same rate as the Karhunen-Loève eigenvalues of the continuous field. Numerical experiments illustrating the theory are given in §4.

2 Circulant Embedding

There are potentially several ways of computing the factorization (1.6). Since R is symmetric positive definite, its Cholesky factorization $R = LL^\top$ yields (1.6), with $s = M$. Alternatively, we can use the spectral decomposition $R = W\Lambda W^\top$, with W being the orthogonal matrix of eigenvectors, and Λ being the positive diagonal matrix of eigenvalues, again giving (1.6) with $s = M$, this time with $B = W\Lambda^{1/2}$. Given that the matrix R is large and dense, in both of these cases the factorization will generally be expensive. In the present paper we use a spectral decomposition of a circulant extension of R to obtain (1.6) with $s > M$ – see Theorem 2.2 below.

2.1 The extended matrix

Recall that we began with a uniform grid of points on the d -dimensional unit cube $[0, 1]^d$. The $M = (m_0 + 1)^d$ points $\{\mathbf{x}_i : i = 1, \dots, M\}$ are assumed to have spacing $h_0 := 1/m_0$ along each coordinate direction. We then consider an enlarged cube $[0, \ell]^d$ of edge length

$$\ell := mh_0 \geq 1, \quad (2.1)$$

with integer $m \geq m_0$. We take the point of view that m_0 is fixed (and hence so too is h_0), while m is variable (and hence so too is ℓ).

We index the grid points on the unit cube by an integer vector \mathbf{k} , writing

$$\mathbf{x}_{\mathbf{k}} := h_0 \mathbf{k} \quad \text{for } \mathbf{k} = (k_1, \dots, k_d) \in \{0, \dots, m_0\}^d. \quad (2.2)$$

Then it is easy to see that (with analogous vector indexing for the rows and columns) the $M \times M$ covariance matrix R defined in (1.4) can be written as

$$R_{\mathbf{k}, \mathbf{k}'} = \rho(h_0(\mathbf{k} - \mathbf{k}')), \quad \mathbf{k}, \mathbf{k}' \in \{0, \dots, m_0\}^d. \quad (2.3)$$

If the vectors \mathbf{k} are enumerated in lexicographical ordering, then R is a nested block Toeplitz matrix where the number of nested levels is the physical dimension d . We remark that all indexing of matrices and vectors by vector notation in this paper is to be considered in this way.

In the following it will be convenient to extend the definition of the grid points (2.2) to an infinite grid,

$$\mathbf{x}_{\mathbf{k}} := h_0 \mathbf{k} \quad \text{for } \mathbf{k} \in \mathbb{Z}^d.$$

Then, in order to define the extended matrix R^{ext} , we define a 2ℓ -periodic map on \mathbb{R} by specifying its action on $[0, 2\ell]$:

$$\varphi(x) := \begin{cases} x & \text{if } 0 \leq x \leq \ell, \\ 2\ell - x & \text{if } \ell \leq x < 2\ell. \end{cases}$$

Now we apply this map componentwise and so define an extended version ρ^{ext} of ρ as follows:

$$\rho^{\text{ext}}(\mathbf{x}) := \rho(\varphi(x_1), \dots, \varphi(x_d)), \quad \mathbf{x} \in \mathbb{R}^d. \quad (2.4)$$

Note that ρ^{ext} is 2ℓ -periodic in each coordinate direction and

$$\rho^{\text{ext}}(\mathbf{x}) = \rho(\mathbf{x}) \quad \text{when } \mathbf{x} \in [0, \ell]^d. \quad (2.5)$$

Then R^{ext} is defined to be the $s \times s$ symmetric nested block circulant matrix with $s = (2m)^d$, defined, analogously to (2.3), by

$$R_{\mathbf{k}, \mathbf{k}'}^{\text{ext}} = \rho^{\text{ext}}(h_0(\mathbf{k} - \mathbf{k}')), \quad \mathbf{k}, \mathbf{k}' \in \{0, \dots, 2m - 1\}^d. \quad (2.6)$$

Moreover, R is the submatrix of R^{ext} in which the indices are constrained to lie in the range $\mathbf{k}, \mathbf{k}' \in \{0, \dots, m_0\}^d$.

In the following it will be convenient to introduce the notation, defined for any integer $m \geq 1$,

$$\mathbb{Z}_{2m}^d := \{0, \dots, 2m - 1\}^d, \quad \overline{\mathbb{Z}}_m^d := \{-m, \dots, m - 1\}^d.$$

Then we have the following simple result.

Proposition 2.1 *R^{ext} has real eigenvalues $\Lambda_{\mathbf{k}}^{\text{ext}}$ and corresponding normalised eigenvectors $\mathbf{V}_{\mathbf{k}}$, given, for $\mathbf{k} \in \mathbb{Z}_{2m}^d$ by the formulae:*

$$\Lambda_{\mathbf{k}}^{\text{ext}} = \sum_{\mathbf{k}' \in \overline{\mathbb{Z}}_m^d} \rho(h_0 \mathbf{k}') \exp\left(-2\pi i \frac{\mathbf{k} \cdot \mathbf{k}'}{2m}\right), \quad \text{and} \quad (\mathbf{V}_{\mathbf{k}})_{\boldsymbol{\kappa}} = \frac{1}{\sqrt{s}} \exp\left(2\pi i \frac{\mathbf{k} \cdot \boldsymbol{\kappa}}{2m}\right), \quad \boldsymbol{\kappa} \in \mathbb{Z}_{2m}^d, \quad (2.7)$$

where s is given in (1.7).

Proof. By (2.4) and the fact that φ is symmetric, we see that ρ^{ext} is also symmetric, so the eigenvalues of R^{ext} are real. To obtain the required formula for the eigenvalues,

we use (2.6) and the formula for the eigenvectors in (2.7) to write, for $\mathbf{k}, \boldsymbol{\kappa} \in \mathbb{Z}_{2m}^d$.

$$\begin{aligned} (R^{\text{ext}} \mathbf{V}_{\mathbf{k}})_{\boldsymbol{\kappa}} &= \sum_{\mathbf{k}' \in \mathbb{Z}_{2m}^d} \rho^{\text{ext}}(h_0(\boldsymbol{\kappa} - \mathbf{k}')) (\mathbf{V}_{\mathbf{k}})_{\mathbf{k}'} = \sum_{\mathbf{k}'' \in \boldsymbol{\kappa} - \mathbb{Z}_{2m}^d} \rho^{\text{ext}}(h_0 \mathbf{k}'') (\mathbf{V}_{\mathbf{k}})_{\boldsymbol{\kappa} - \mathbf{k}''} \\ &= \left(\sum_{\mathbf{k}'' \in \boldsymbol{\kappa} - \mathbb{Z}_{2m}^d} \rho^{\text{ext}}(h_0 \mathbf{k}'') \exp \left(-2\pi i \frac{\mathbf{k} \cdot \mathbf{k}''}{2m} \right) \right) (\mathbf{V}_{\mathbf{k}})_{\boldsymbol{\kappa}}. \end{aligned}$$

Then (2.7) follows, on using the coordinatewise $2m$ -periodicity of the summand in the last equation, and also the extension property (2.5). \square

It follows from the simple form of the eigenfunctions that the matrix R^{ext} can be diagonalised by FFT. The following version of the spectral decomposition theorem, taken from [11], has the advantage that it allows the diagonalisation to be implemented using only real FFT.

Theorem 2.2 *R^{ext} has the spectral decomposition:*

$$R^{\text{ext}} = Q^{\text{ext}} \Lambda^{\text{ext}} Q^{\text{ext}},$$

where Λ^{ext} is the diagonal matrix containing the eigenvalues of R^{ext} , and $Q^{\text{ext}} = \Re(\mathcal{F}) + \Im(\mathcal{F})$ is real symmetric, with

$$\mathcal{F}_{\mathbf{k}, \mathbf{k}'} = \frac{1}{\sqrt{s}} \exp \left(2\pi i \frac{\mathbf{k}' \cdot \mathbf{k}}{2m} \right)$$

denoting the d -dimensional Fourier matrix. If the eigenvalues of R^{ext} are all non-negative then the required B in (1.6) can be obtained by selecting M appropriate rows of

$$B^{\text{ext}} := Q^{\text{ext}} (\Lambda^{\text{ext}})^{1/2}. \quad (2.8)$$

The use of FFT allows fast computation of the matrix-vector product $B^{\text{ext}} \mathbf{y}$ for any vector \mathbf{y} , which then yields $B \mathbf{y}$ needed for sampling the random field in (1.5). Our algorithm for obtaining a minimal positive definite R^{ext} is given in Algorithm 1. Our algorithm for sampling an instance of the random field is given in Algorithm 2. Note that the normalisation used within the FFT routine differs among particular implementations. Here, we assume the Fourier transform to be unitary.

We can replace Step 1 of Algorithm 2 with a QMC point from $[0, 1]^s$ and mapped to \mathbb{R}^s elementwise by the inverse of the cumulative normal distribution function. The relative size of the eigenvalues in Λ^{ext} tells us the relative importance of the corresponding variables in the extended system, which helps to determine the ordering of the QMC variables.

Algorithm 1 *Input: d , m_0 , and covariance function ρ .*

1. Set $m = m_0$.
2. Calculate \mathbf{r} , the first column of R^{ext} in (2.6).
3. Calculate \mathbf{v} , the vector of eigenvalues of R^{ext} , by d -dimensional FFT on \mathbf{r} .

4. If smallest eigenvalue < 0 then increment m and go to Step 2.

Output: m, \mathbf{v} .

Algorithm 2 Input: d, m_0 , mean field \bar{Z} , and m and \mathbf{v} obtained by Algorithm 1.

1. With $s = (2m)^d$, sample an s -dimensional normal random vector \mathbf{y} .
2. Update \mathbf{y} by elementwise multiplication with $\sqrt{\mathbf{v}}$.
3. Set \mathbf{w} to be the d -dimensional FFT of \mathbf{y} .
4. Update \mathbf{w} by adding its real and imaginary parts.
5. Obtain \mathbf{z} by extracting the appropriate $M = (m_0 + 1)^d$ entries of \mathbf{w} .
6. Update \mathbf{z} by adding \bar{Z} .

Output: \mathbf{z} (or $\exp(\mathbf{z})$ in the case of lognormal field).

In the following subsection we shall show (under mild conditions) that Algorithm 1 will always terminate. Moreover we shall give (for the case of the Matérn covariance function) a detailed analysis of how ℓ depends on various parameters of the field. Then in §3, we give an analysis of the decay rates of the eigenvalues of R^{ext} compared with that of the eigenvalues of the original Toeplitz matrix R and the KL eigenvalues of the underlying continuous field Z .

2.2 Positive definiteness

We first note that by definition (1.1) and (1.2) it follows that for all $N \geq 1$, all point sets $\mathbf{t}_1, \dots, \mathbf{t}_N \in \mathbb{R}^d$, and all $\boldsymbol{\gamma} \in \mathbb{R}^N$

$$\sum_{i=1}^N \sum_{i'=1}^N \gamma_i \gamma_{i'} \rho(\mathbf{t}_i - \mathbf{t}_{i'}) = \mathbb{E}[(W - \bar{W})^2] \geq 0, \quad (2.9)$$

where

$$W(\omega) = \sum_{i=1}^N \gamma_i Z(\mathbf{t}_i, \omega)$$

and \bar{W} denotes its mean. This observation, together with (1.3), means that ρ is a *symmetric positive semidefinite* function (see, e.g., [22, Chapter 6]). If in addition, (2.9) is always positive for nonzero $\boldsymbol{\gamma}$ then ρ is called *positive definite*. In our later applications, ρ will always be positive definite.

We use the following definition of the Fourier transform of an absolutely integrable function ρ :

$$\hat{\rho}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} \rho(\mathbf{x}) \exp(-2\pi i \boldsymbol{\xi} \cdot \mathbf{x}) d\mathbf{x}, \quad \boldsymbol{\xi} \in \mathbb{R}^d. \quad (2.10)$$

If, in addition, ρ is continuous and $\widehat{\rho}$ is absolutely integrable then the Fourier integral theorem gives

$$\rho(\mathbf{x}) = \int_{\mathbb{R}^d} \widehat{\rho}(\boldsymbol{\xi}) \exp(2\pi i \boldsymbol{\xi} \cdot \mathbf{x}) d\boldsymbol{\xi}, \quad \mathbf{x} \in \mathbb{R}^d.$$

In this case, $\widehat{\rho}$ is also continuous, and Bochner's theorem (e.g., [18, Theorem 6.3]) states that ρ is a positive definite function if and only if $\widehat{\rho}$ is positive.

The following theorem is the main result of this subsection.

Theorem 2.3 *Suppose that $\rho \in L^1(\mathbb{R}^d)$ is a real-valued, symmetric positive definite function with the additional reflectional symmetry*

$$\rho(\mathbf{x}) = \rho(\pm x_1, \dots, \pm x_d) = \rho(|x_1|, \dots, |x_d|) \quad \text{for all } \mathbf{x} \in \mathbb{R}^d,$$

and suppose $\widehat{\rho} \in L^1(\mathbb{R}^d)$. Suppose also that for the given value of $h_0 = 1/m_0$ we have

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} |\rho(h_0 \mathbf{k})| < \infty. \quad (2.11)$$

Then Algorithm 1 will always terminate with a finite value of m , and the resulting matrix R^{ext} will be positive definite.

We prove Theorem 2.3 via a technical estimate – Lemma 2.4 – which provides an explicit lower bound for the eigenvalues of R^{ext} . This estimate proves the theorem, and moreover, allows us to obtain explicit lower bounds for the eigenvalues in the lemmas which follow.

Lemma 2.4 *Under the assumptions of Theorem 2.3, the eigenvalues $\Lambda_{\mathbf{k}}^{\text{ext}}$ of R^{ext} all satisfy the estimate*

$$\Lambda_{\mathbf{k}}^{\text{ext}} \geq \frac{1}{h_0^d} \min_{\boldsymbol{\zeta} \in [-\frac{1}{2}, \frac{1}{2}]^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \widehat{\rho}\left(\frac{\boldsymbol{\zeta} + \mathbf{r}}{h_0}\right) - \sum_{\mathbf{k}' \in \mathbb{Z}^d \setminus \overline{\mathbb{Z}}_m^d} |\rho(h_0 \mathbf{k}')|. \quad (2.12)$$

Proof. By Proposition 2.1, we have

$$\Lambda_{\mathbf{k}}^{\text{ext}} = \sum_{\mathbf{k}' \in \mathbb{Z}^d} \rho(h_0 \mathbf{k}') \exp\left(-2\pi i \frac{\mathbf{k}' \cdot \mathbf{k}}{2m}\right) - \sum_{\mathbf{k}' \in \mathbb{Z}^d \setminus \overline{\mathbb{Z}}_m^d} \rho(h_0 \mathbf{k}') \exp\left(-2\pi i \frac{\mathbf{k}' \cdot \mathbf{k}}{2m}\right). \quad (2.13)$$

We obtain the lower bound on the first sum on the right-hand side of (2.13) using (A.3) of Theorem A.1 in the Appendix with $h = h_0$ and $\boldsymbol{\xi} = \mathbf{k}'/(2m)$, and an upper bound on the second sum in the obvious way. \square

Proof of Theorem 2.3. Because of the assumed positive definiteness of ρ , it follows from Bochner's theorem that the Fourier transform $\widehat{\rho}$ is positive, so the strict positivity of the first term on the right-hand side of (2.12) follows from Theorem A.1. The lower bound is also independent of m . For fixed h_0 , (2.11) ensures that the tail sum in the second term on the right-hand side of (2.12) converges to zero as $m \rightarrow \infty$. Hence the result follows. \square

2.3 Isotropic covariance

More detailed lower bounds on $\Lambda_{\mathbf{k}}^{\text{ext}}$ can be obtained under the additional assumption that the random field Z in (1.1) is *isotropic*, i.e.,

$$\rho(\mathbf{x}) = \kappa(\|\mathbf{x}\|_2/\lambda), \quad (2.14)$$

where the parameter λ is a correlation length which will play a key role in Example 2.7 and Theorem 2.9 below. In this case the Fourier transform (2.10) is given by

$$\widehat{\rho}(\boldsymbol{\xi}) = \lambda^d \widehat{\kappa}_d(\lambda \|\boldsymbol{\xi}\|_2), \quad (2.15)$$

where

$$\widehat{\kappa}_d(r) := \frac{2\pi}{r^{(d-2)/2}} \int_0^\infty \kappa(t) t^{d/2} J_{(d-2)/2}(2\pi r t) dt \quad \text{with } r \geq 0, \quad (2.16)$$

and J_α denotes the Bessel function of order α . The right-hand side of (2.16) is the *Hankel transform* of κ (see, e.g., [18, Theorem 1.107]). The following lemma then estimates each of the terms on the right-hand side of (2.12) to obtain an explicit lower bound on the eigenvalues of R^{ext} .

Lemma 2.5 *Suppose ρ and $\widehat{\rho}$ are given as in (2.14)–(2.16).*

- (i) *If $|\kappa|$ is a decreasing function on \mathbb{R}_+ and $r^{d-1} \kappa(r) \in L_1(\mathbb{R}_+)$, then (with $\ell = mh_0$),*

$$\sum_{\mathbf{k} \in \mathbb{Z}^d \setminus \overline{\mathbb{Z}}_m^d} |\rho(h_0 \mathbf{k})| \leq \frac{(3^d - 1)2^{d-1}}{(h_0/\lambda)^d} \int_{(\ell-h_0)/\lambda}^\infty r^{d-1} |\kappa(r)| dr.$$

- (ii) *If $\widehat{\kappa}_d$ is positive and decreasing on \mathbb{R}_+ and $r^{d-1} \widehat{\kappa}_d(r) \in L_1(\mathbb{R}_+)$, then*

$$\frac{1}{h_0^d} \min_{\boldsymbol{\zeta} \in [-\frac{1}{2}, \frac{1}{2}]^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \widehat{\rho}\left(\frac{\boldsymbol{\zeta} + \mathbf{r}}{h_0}\right) \geq \frac{2^d}{d^{d/2-1} 3^{d-1}} \int_{3d^{1/2}\lambda/(2h_0)}^\infty r^{d-1} \widehat{\kappa}_d(r) dr.$$

Proof. We begin by deriving upper and lower bounds for some elementary sequences.

Suppose g is any positive and decreasing function satisfying $r^{d-1}g(r) \in L_1(\mathbb{R}_+)$. For any integer $m \geq 1$ we have

$$\sum_{j=m+1}^\infty j^{d-1} g(j) \leq \sum_{j=m+1}^\infty \int_{j-1}^j [r]^{d-1} g(r) dr \leq \int_m^\infty (2r)^{d-1} g(r) dr, \quad (2.17)$$

where the first inequality uses $[r] = j$ for $r \in (j-1, j]$ and the second inequality follows from $[r] \leq 2r$ for $r \geq 1$. Similarly, we can obtain a lower bound for $m \geq 1$,

$$\sum_{j=m+1}^\infty j^{d-1} g(j) \geq \sum_{j=m+1}^\infty \int_j^{j+1} [r]^{d-1} g(r) dr \geq \int_{m+1}^\infty (r/2)^{d-1} g(r) dr. \quad (2.18)$$

Also let us consider the set $S_d(j) := \{\mathbf{k} \in \mathbb{Z}^d : \|\mathbf{k}\|_\infty = j\}$ for any integer $j \geq 1$. Then, with $\# S_d(j)$ denoting the cardinality of this set, we have the bounds

$$d 2^d j^{d-1} \leq \# S_d(j) = (2j+1)^d - (2j-1)^d = 2 \sum_{\substack{i=1 \\ i \text{ odd}}}^d \binom{d}{i} (2j)^{d-i} \leq (3^d - 1) j^{d-1}, \quad (2.19)$$

where the lower bound is the $i = 1$ term in the binomial expansion and the upper bound comes from the estimate $j^{d-i} \leq j^{d-1}$ for $i \geq 1$. The bounds are exact for $d = 1$ and $d = 2$, with $\# S_1(j) = 2$ and $\# S_2(j) = 8j$, while for $d = 3$ we have $\# S_3(j) = 2 + 24j^2$, and the lower and upper bounds given by (2.19) are $24j^2$ and $26j^2$ respectively.

Now, to prove (i), using (2.19) and (2.17) and $\|\mathbf{k}\|_2 \geq \|\mathbf{k}\|_\infty$, we can now write

$$\begin{aligned} \sum_{\mathbf{k} \in \mathbb{Z}^d \setminus \bar{\mathbb{Z}}_m^d} |\kappa(h_0 \|\mathbf{k}\|_2 / \lambda)| &\leq \sum_{j=m}^{\infty} \sum_{\|\mathbf{k}\|_\infty=j} |\kappa(h_0 \|\mathbf{k}\|_2 / \lambda)| \leq (3^d - 1) \sum_{j=m}^{\infty} j^{d-1} |\kappa(h_0 j / \lambda)| \\ &\leq (3^d - 1) 2^{d-1} \int_{m-1}^{\infty} r^{d-1} |\kappa(h_0 r / \lambda)| dr = \frac{(3^d - 1) 2^{d-1}}{(h_0 / \lambda)^d} \int_{(\ell-h_0)/\lambda}^{\infty} r^{d-1} |\kappa(r)| dr, \end{aligned}$$

with $\ell = m h_0$, thus completing the proof of (i).

To prove (ii), note first that, for $\boldsymbol{\zeta} \in [-\frac{1}{2}, \frac{1}{2}]^d$ and any $\mathbf{r} \in \mathbb{Z}^d$ with $\|\mathbf{r}\|_\infty = j$, we have $\|\boldsymbol{\zeta} + \mathbf{r}\|_2 \leq d^{1/2} \|\boldsymbol{\zeta} + \mathbf{r}\|_\infty \leq d^{1/2} (1/2 + j)$. Using (2.15), (2.19) and (2.18) and dropping the $\mathbf{r} = \mathbf{0}$ term, we can write

$$\begin{aligned} \frac{1}{h_0^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \hat{\rho}\left(\frac{\boldsymbol{\zeta} + \mathbf{r}}{h_0}\right) &= \left(\frac{\lambda}{h_0}\right)^d \sum_{\mathbf{r} \in \mathbb{Z}^d} \hat{\kappa}_d\left(\frac{\lambda \|\boldsymbol{\zeta} + \mathbf{r}\|_2}{h_0}\right) \geq \left(\frac{\lambda}{h_0}\right)^d \sum_{j=1}^{\infty} \sum_{\|\mathbf{r}\|_\infty=j} \hat{\kappa}_d\left(\frac{\lambda \|\boldsymbol{\zeta} + \mathbf{r}\|_2}{h_0}\right) \\ &\geq \frac{d 2^d}{(h_0 / \lambda)^d} \sum_{j=1}^{\infty} j^{d-1} \hat{\kappa}_d\left(\lambda \frac{d^{1/2}(1/2 + j)}{h_0}\right) \geq \frac{2d}{(h_0 / \lambda)^d} \int_1^{\infty} r^{d-1} \hat{\kappa}_d\left(\frac{\lambda d^{1/2}(1/2 + r)}{h_0}\right) dr \\ &= \frac{2}{d^{d/2-1}} \int_{3d^{1/2}\lambda/(2h_0)}^{\infty} \left(r - \frac{\lambda d^{1/2}}{2h_0}\right)^{d-1} \hat{\kappa}_d(r) dr \geq \frac{2}{d^{d/2-1}} \int_{3d^{1/2}\lambda/(2h_0)}^{\infty} \left(\frac{2r}{3}\right)^{d-1} \hat{\kappa}_d(r) dr, \end{aligned}$$

where the last inequality follows from $r - c \geq 2r/3 \Leftrightarrow r \geq 3c$, with $c = d^{1/2}\lambda/(2h_0)$. \square

Corollary 2.6 *Under the assumptions of Lemma 2.5, R^{ext} is positive definite if*

$$\int_{3\lambda d^{1/2}/(2h_0)}^{\infty} r^{d-1} \hat{\kappa}_d(r) dr > \frac{(3^d - 1) 3^{d-1} d^{d/2-1}}{2(h_0 / \lambda)^d} \int_{(\ell-h_0)/\lambda}^{\infty} r^{d-1} |\kappa(r)| dr. \quad (2.20)$$

Proof. We make use of Lemma 2.4. The fact that $\rho \in L_1(\mathbb{R}^d)$ and $\hat{\rho} \in L_1(\mathbb{R}^d)$ follow immediately from (2.14) and (2.16) respectively, and from the assumptions $r^{d-1} \kappa(r) \in L_1(\mathbb{R}^+)$ and $r^{d-1} \hat{\kappa}_d(r) \in L_1(\mathbb{R}^+)$. It then follows from Part (i) of Lemma 2.5 that the assumption (2.11) of Theorem 2.3 is satisfied. Since the symmetry assumption in the theorem is automatically satisfied by an isotropic covariance, the result now follows immediately from Lemma 2.4. \square

To interpret Corollary 2.6, recall that λ is the correlation length, so λ is bounded above (without loss of generality let us assume $\lambda \leq 1$), but λ may approach 0. If h_0 is chosen so that h_0/λ is a fixed constant, then (2.20) and integrability of $r^{d-1}|\kappa(r)|$ ensures that positive definiteness is achieved for ℓ sufficiently large independently of λ and h_0 . The condition “ h_0/λ constant” is a natural analogue to the requirement in oscillatory problems that the meshwidth should be proportional to the wavelength. However if ℓ and λ are fixed and $h_0 \rightarrow 0$ (constituting “subwavelength mesh refinement” needed to get higher accuracy) we see that the sufficient condition that ensures positive definiteness will eventually fail. One interesting question is how fast ℓ needs to grow as h_0 decreases in order to be sure of positive definiteness. This can be answered fairly completely and satisfactorily in the case of the Matérn family.

Example 2.7 The Matérn family of covariances are defined by

$$\rho(\mathbf{x}) = \kappa(\|\mathbf{x}\|_2/\lambda) , \quad \text{where} \quad \kappa(r) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (\sqrt{2\nu} r)^\nu K_\nu(\sqrt{2\nu} r) . \quad (2.21)$$

Here Γ is the gamma function and K_ν is the modified Bessel function of the second kind, σ^2 is the variance, λ is the correlation length and $\nu > 0$ is a smoothness parameter. The practically interesting range is $\nu \geq 1/2$, with the cases $\nu = 1/2$ and $\nu = \infty$ corresponding to the exponential and Gaussian covariances respectively, see, e.g., [13]. For this reason, we have restricted the analysis to the case $\nu \geq 1/2$, see also Remark 2.10.

For this particular κ , the Hankel transform $\widehat{\kappa}_d$ (see (2.16)) is explicitly known, and so we have from (2.15)–(2.16) that

$$\widehat{\rho}(\boldsymbol{\xi}) = \lambda^d \widehat{\kappa}_d(\lambda \|\boldsymbol{\xi}\|_2) , \quad \text{where} \quad \widehat{\kappa}_d(r) = \sigma^2 2^d \pi^{d/2} (2\nu)^\nu \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \frac{1}{(2\nu + (2\pi r)^2)^{\nu+d/2}} . \quad (2.22)$$

(This can be obtained, for example, by some manipulation of the formula [18, p.264].) Note that for the Matérn case both ρ and $\widehat{\rho}$ are positive, radial, decreasing functions.

The following theorem shows that when ν and λ are fixed, ℓ needs to grow with order $\log h_0^{-1}$ in order to be sure of positive definiteness as $h_0 \rightarrow 0$. On the other hand if λ and h_0 are fixed, ℓ needs to grow like $\nu^{1/2} \log \nu$ as ν increases. For given ℓ , h_0 , the bound on ℓ gets smaller as λ decreases. This provides conditions guaranteeing the termination of Algorithm 1.

Notation 2.8 When discussing the Matérn case, we shall use the notation $A \lesssim B$ (equivalently $B \gtrsim A$) to mean A/B is bounded above independently of ℓ, h_0, λ and ν , and we write $A \sim B$ if $A \lesssim B$ and $B \lesssim A$.

Theorem 2.9 Consider the Matérn covariance family (2.21), with smoothness parameter ν satisfying $1/2 \leq \nu < \infty$ and correlation length $\lambda \leq 1$. Suppose $h_0/\lambda \leq e^{-1}$. Then there exist positive constants C_1 and $C_2 \geq 2\sqrt{2}$ which may depend on dimension d but are independent of the other parameters $\ell, h_0, \lambda, \nu, \sigma^2$, such that R^{ext} is positive definite if

$$\ell/\lambda \geq C_1 + C_2 \nu^{1/2} \log(\max\{\lambda/h_0, \nu^{1/2}\}) . \quad (2.23)$$

Proof. For convenience we introduce the notation

$$\Psi = \max\{\lambda/h_0, \nu^{1/2}\}.$$

Aiming to verify (2.20), we note first that, since both κ and $\hat{\kappa}_d$ depend linearly on σ^2 , we can without loss of generality set $\sigma^2 = 1$. We shall obtain the following lower bound on the left-hand side of (2.20):

$$\int_{3\lambda d^{1/2}/(2h_0)}^{\infty} r^{d-1} \hat{\kappa}_d(r) dr \gtrsim \nu^{\nu+d/2-1} (4\pi^2)^{-\nu} \Psi^{-2\nu}, \quad (2.24)$$

and (subject to assumption (2.23)), the following upper bound on the integral on the right-hand side of (2.20):

$$\int_{(\ell-h_0)/\lambda}^{\infty} r^{d-1} |\kappa(r)| dr \lesssim 5^\nu \nu^{d/2-1} \exp\left(-\sqrt{\frac{\nu}{2}} \frac{\ell-h_0}{\lambda}\right). \quad (2.25)$$

Since these estimates are rather technical, we defer their justification until the end of this proof. Thus, assuming (2.24) and (2.25) we see that there exists $D > 0$, with D independent of ℓ, h_0, λ, ν , such that (2.20) holds if

$$\exp\left(-\sqrt{\frac{\nu}{2}} \frac{\ell-h_0}{\lambda}\right) \leq D \left(\frac{\nu}{20\pi^2}\right)^\nu \Psi^{-2\nu}.$$

Taking logs and rearranging, this is equivalent to

$$(\ell-h_0)/\lambda \geq \sqrt{2/\nu} \log(1/D) + \sqrt{2}\nu^{1/2} \log(20\pi^2/\nu) + 2\sqrt{2}\nu^{1/2} \log \Psi.$$

The first term on the right-hand side is positive only when $D < 1$, in which case it is sufficient to replace $\sqrt{2/\nu}$ by its maximum value 2 (given that $\nu \geq 1/2$). Similarly, the second term is positive only when $\nu \leq 20\pi^2$, and in this case we have

$$\nu^{1/2} \log(20\pi^2/\nu) \leq \sqrt{20\pi^2} \log(40\pi^2).$$

Taking into account that $h_0/\lambda \leq e^{-1}$, we have thus demonstrated the sufficiency of a condition of the form (2.23). We now complete the proof by proving the technical estimates (2.24) and (2.25).

Proof of estimate (2.24): Recalling that $\hat{\kappa}_d$ is given by (2.22), we may use the following elementary result to bound the ratio of gamma functions if d is even:

$$\frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} = (\nu + d/2 - 1) \dots (\nu + 1) \nu \geq \nu^{d/2}. \quad (2.26)$$

If d is odd then we may use

$$\begin{aligned} \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} &= (\nu + d/2 - 1) \dots (\nu + 3/2)(\nu + 1/2) \frac{\Gamma(\nu + 1/2)}{\Gamma(\nu)} \\ &\geq \nu^{(d-1)/2} \frac{\Gamma(\nu + 1/2)}{\Gamma(\nu)} \geq \nu^{(d-1)/2} (\nu - 1/2 + 1/4)^{1/2} \geq \nu^{(d-1)/2} \nu^{1/2} / \sqrt{2} \end{aligned} \quad (2.27)$$

for all $\nu \geq 1/2$, where in the penultimate step we use Kershaw's inequality (see [15], equation (1.3))

$$\frac{\Gamma(x+1)}{\Gamma(x+r)} > \left(x + \frac{r}{2}\right)^{1-r} \quad \text{for } x > 0, \quad 0 < r < 1,$$

with $r = 1/2$ and $x = \nu - 1/2$. (If $\nu = 1/2$ the result is obtained by taking the limit $\nu \rightarrow 1/2+$ and using continuity.)

On using this lower bound in (2.22), we obtain

$$\widehat{\kappa}_d(r) \gtrsim (2\nu)^\nu \nu^{d/2} (2\nu)^{-(\nu+d/2)} (1 + 2\pi^2 r^2/\nu)^{-(\nu+d/2)} \sim (1 + 2\pi^2 r^2/\nu)^{-(\nu+d/2)}.$$

Now (for convenience), we introduce the notation

$$a = 3\sqrt{d}/2, \quad \widetilde{\Psi} = \max\{\lambda a/h_0, \sqrt{\nu/2\pi^2}\}.$$

Then the left-hand side of (2.20) can be estimated from below by:

$$\begin{aligned} \int_{\lambda a/h_0}^{\infty} r^{d-1} \widehat{\kappa}_d(r) \, dr &\gtrsim \int_{\lambda a/h_0}^{\infty} r^{d-1} (1 + 2\pi^2 r^2/\nu)^{-(\nu+d/2)} \, dr \\ &\geq \int_{\widetilde{\Psi}}^{\infty} r^{d-1} (1 + 2\pi^2 r^2/\nu)^{-(\nu+d/2)} \, dr. \end{aligned}$$

Now, noting that when $r \geq \widetilde{\Psi}$ we have $2\pi^2 r^2/\nu \geq 1$, and so

$$\begin{aligned} \int_{\lambda a/h_0}^{\infty} r^{d-1} \widehat{\kappa}_d(r) \, dr &\geq \int_{\widetilde{\Psi}}^{\infty} r^{d-1} (4\pi^2 r^2/\nu)^{-(\nu+d/2)} \, dr \\ &= \left(\frac{\nu}{4\pi^2}\right)^{\nu+d/2} \int_{\widetilde{\Psi}}^{\infty} r^{-2\nu-1} \, dr \sim \nu^{\nu+d/2-1} (4\pi^2)^{-(\nu+d/2)} \widetilde{\Psi}^{-2\nu}. \end{aligned}$$

which yields (2.24), since $\widetilde{\Psi} \sim \Psi$.

Proof of estimate (2.25): It is sufficient to prove this estimate with $\ell - h_0$ on each side replaced by ℓ . By definition (2.21) and the change of variable $r \mapsto (\sqrt{2/\nu})r$, we have

$$\int_{\ell/\lambda}^{\infty} r^{d-1} |\kappa(r)| \, dr \sim \frac{1}{\Gamma(\nu)} \left(\frac{\nu}{2}\right)^{\nu+d/2} \int_{\sqrt{2/\nu}(\ell/\lambda)}^{\infty} r^{d+\nu-1} |K_\nu(\nu r)| \, dr. \quad (2.28)$$

To obtain an estimate for the Bessel function on the right-hand side of (2.28), note first that by the hypotheses of this theorem, we have $h_0/\lambda \leq e^{-1}$ and $\nu \geq 1/2$. Now assume that (2.23) holds, with $C_1 > 0$ and $C_2 > 2\sqrt{2}$, both not yet fixed. Then

$$\ell/\lambda \geq C_1 + C_2 \nu^{1/2} \geq C_1 + 2\sqrt{2} \nu^{1/2} > 2\sqrt{2} \nu^{1/2} = 4\sqrt{\nu/2}, \quad (2.29)$$

and hence the range of integration in (2.28) is contained in $[4, \infty)$.

The uniform asymptotic estimate for $K_\nu(\nu r)$ given in [1, 9.7.8] implies that there exists a $\nu^* < \infty$ and a constant C such that, for all $\nu \geq \nu^*$,

$$|K_\nu(\nu r)| \leq C \nu^{-1/2} \exp(-\nu r) \quad r \in [4, \infty). \quad (2.30)$$

We shall show that in fact such an inequality holds for all $\nu \geq 1/2$ and all $r \geq 4$. To this end, for $\nu \geq 1/2$, we introduce the quantity

$$c(\nu) := \|\exp(z)z^{1/2}K_\nu(z)\|_{L_\infty(2,\infty)},$$

The continuity of K_ν on $[2, \infty)$ and the order-dependent asymptotics of K_ν [1, 9.7.2] ensure that $c(\nu) < \infty$ for all $\nu < \infty$. It can also be shown, by appealing to the integral representation of the modified Bessel function, that $c(\nu)$ is continuous with respect to ν when $\nu \geq 1/2$, from which we deduce that $c^* := \max_{[1/2, \nu^*]} c(\nu) < \infty$. Now, for $1/2 \leq \nu \leq \nu^*$ and $r \geq 4$ we have $\nu r \geq 2$ and

$$\exp(\nu r)\nu^{1/2}|K_\nu(\nu r)| = r^{-1/2} [\exp(\nu r)(\nu r)^{1/2}|K_\nu(\nu r)|] \leq \frac{1}{2}c(\nu) \leq \frac{1}{2}c^*.$$

This shows that the estimate (2.30) holds uniformly for all $r \geq 4$ and $\nu \geq 1/2$. Using this in (2.28) we then have

$$\int_{\ell/\lambda}^{\infty} r^{d-1}|\kappa(r)|dr \lesssim \frac{1}{\Gamma(\nu)} \left(\frac{\nu}{2}\right)^{\nu+d/2} \nu^{-1/2} \int_{\sqrt{2/\nu}(\ell/\lambda)}^{\infty} r^{d-1+\nu} e^{-\nu r} dr. \quad (2.31)$$

Now note that (2.29) implies that $\ell/\lambda \geq C_2\nu^{1/2}$, and hence $\sqrt{2/\nu}(\ell/\lambda) \geq \sqrt{2}C_2$, so we can choose C_2 a large enough constant independent of $\ell, h_0, \lambda, \nu, \sigma^2$ so that $\sqrt{2/\nu}(\ell/\lambda) \geq 10$. Then, using the elementary inequality $xe \leq e^x$ which holds when $x \geq 1$, it follows that $(r/10)e \leq \exp(r/10)$ when $r \geq 10$. Hence for $r \geq \sqrt{2/\nu}(\ell/\lambda) \geq 10$, we have

$$r^{d-1+\nu} \leq \left(\frac{10}{e}\right)^{d-1+\nu} \exp\left(\frac{d-1+\nu}{10}r\right) \leq \left(\frac{10}{e}\right)^{d-1+\nu} e^{\nu r/2},$$

where in the last step we used $d-1 \leq 2 \leq 4\nu$. Inserting this into the right-hand side of (2.31), we get, after integration and some manipulation,

$$\int_{\ell/\lambda}^{\infty} r^{d-1}|\kappa(r)|dr \lesssim \frac{1}{\Gamma(\nu)} \left(\frac{5\nu}{e}\right)^{\nu} \nu^{d/2-3/2} \exp\left(-\sqrt{\frac{\nu}{2}}\frac{\ell}{\lambda}\right).$$

Stirling's formula implies that $\Gamma(\nu)\nu^{1/2}e^{\nu} \sim \nu^{\nu}$ and the estimate (2.25) follows. \square

Remark 2.10 The cases $\nu \geq 1/2$ in the Matérn family of covariances are of main interest in applications. In order to avoid further technicalities, we have restricted our attention to those cases in the proof and used the lower bound on ν several times in a non-trivial way. First of all it is used in the demonstration that estimates (2.23), (2.24) and (2.25) are sufficient to ensure the positive definiteness of R^{ext} . Then it is used in the proofs of each of the estimates (2.24) and (2.25). The extension of Theorem 2.9 to $\nu \in (0, 1/2)$ remains an open question.

To complete the theory of this section, we discuss the case $\nu = \infty$. In this case,

$$\rho(\mathbf{x}) = \kappa(\|\mathbf{x}\|_2/\lambda), \quad \text{with} \quad \kappa(r) = \sigma^2 \exp(-r^2/2),$$

and an elementary calculation gives

$$\widehat{\rho}(\boldsymbol{\xi}) = \sigma^2 (2\pi)^{d/2} \lambda^d \exp(-2\pi^2 \lambda^2 \|\boldsymbol{\xi}\|_2^2).$$

For this (Gaussian) covariance it is well-known that the Karhunen-Lo  ve expansion converges exponentially (see, e.g., [21]), in which case it may be preferable to compute realisations of the field Z via the KL expansion, rather than the method proposed here. Nevertheless the existing analysis can be applied to this case as we now show.

Theorem 2.11 *For the Gaussian covariance (i.e., the Mat  rn kernel with $\nu = \infty$), there exists a constant B (depending only on spatial dimension d) such that positive definiteness of R^{ext} is guaranteed when*

$$\ell \geq 1 + \lambda \max \left\{ \sqrt{2} \frac{\lambda}{h_0}, B \right\}.$$

Hence, if λ/h_0 is fixed (i.e., a fixed number of grid points per unit correlation length is used), then the ℓ required for positive definiteness decreases linearly with decreasing correlation length λ , until the minimal value $\ell = 1$ is reached.

Proof. As before, without loss of generality we set $\sigma^2 = 1$ and we make use of Corollary 2.6. The left-hand side of (2.20) is then

$$(2\pi)^{d/2} \int_{3\lambda d^{1/2}/(2h_0)}^{\infty} r^{d-1} \exp(-2\pi^2 r^2) dr = (2\pi)^{-d/2} \int_{c_1 \lambda/h_0}^{\infty} r^{d-1} \exp(-r^2/2) dr,$$

with $c_1 = 3\pi\sqrt{d} > 1$. To make this quantity greater than the right-hand side of (2.20), we require

$$\int_{(\ell-h_0)/\lambda}^{\infty} r^{d-1} \exp(-r^2/2) dr < c_2 \left(\frac{h_0}{\lambda} \right)^d \int_{c_1 \lambda/h_0}^{\infty} r^{d-1} \exp(-r^2/2) dr,$$

with $c_2 = 2d(2\pi d)^{-d/2} (3^d - 1)^{-1} 3^{1-d}$. We shall show that there exists B (depending only on d) such that

$$\int_{\ell/\lambda}^{\infty} r^{d-1} \exp(-r^2/2) dr < c_2 \left(\frac{h_0}{\lambda} \right)^d \int_{c_1 \lambda/h_0}^{\infty} r^{d-1} \exp(-r^2/2) dr, \quad (2.32)$$

when

$$\ell \geq \lambda \max \left\{ \sqrt{2} c_1 \frac{\lambda}{h_0}, B \right\}, \quad (2.33)$$

and the statement of the theorem then follows, since $h_0 \leq 1$.

To prove (2.32), we estimate its left-hand side by

$$\begin{aligned} \int_{\ell/\lambda}^{\infty} r^{d-1} \exp(-r^2/2) dr &\leq \exp(-\ell^2/4\lambda^2) \int_{\ell/\lambda}^{\infty} r^{d-1} \exp(-r^2/4) dr \\ &= 2^{d/2} \exp(-\ell^2/4\lambda^2) \int_{\ell/(\sqrt{2}\lambda)}^{\infty} r^{d-1} \exp(-r^2/2) dr. \end{aligned} \quad (2.34)$$

We now choose B (depending only on d) to have the property that

$$\exp(-x^2/4) x^d \leq c_2, \quad \text{when } x \geq B. \quad (2.35)$$

Then if ℓ satisfies (2.33), we have

$$\ell/\lambda \geq B \quad \text{and} \quad \ell/\lambda \geq \sqrt{2}c_1\lambda/h_0 > \sqrt{2}\lambda/h_0. \quad (2.36)$$

Thus, using (2.35) and (2.36), we have

$$2^{d/2} \exp(-\ell^2/4\lambda^2) \leq c_2 \left(\sqrt{2}\lambda/\ell \right)^d < c_2 (h_0/\lambda)^d.$$

Combining this with (2.34) and using the fact that $\ell/(\sqrt{2}\lambda) \geq c_1\lambda/h_0$ we obtain (2.32). \square

3 Eigenvalue decay

In this section we return to the formula (1.5) for sampling the random field \mathbf{Z} . For the circulant embedding method, using the notation introduced above, we have

$$\mathbf{Z}(\omega) = \sum_{\mathbf{k} \in \overline{\mathbb{Z}}_m^d} \mathbf{B}_{\mathbf{k}} Y_{\mathbf{k}}(\omega) + \overline{\mathbf{Z}},$$

where $Y_{\mathbf{k}}$ are i.i.d. standard normals, $\mathbf{B}_{\mathbf{k}}$ are the columns of the matrix B , and in this case B is taken to be the appropriate M rows of the matrix B^{ext} , as described in Theorem 2.2. Recalling (2.8) and noting that every entry in $\Re(\mathcal{F}) + \Im(\mathcal{F})$ is bounded by $\sqrt{2/s}$, it follows that

$$\|\mathbf{B}_{\mathbf{k}}\|_{\infty} \leq \sqrt{\frac{2\Lambda_{\mathbf{k}}^{\text{ext}}}{s}}, \quad \mathbf{k} \in \overline{\mathbb{Z}}_m^d,$$

where $\Lambda_{\mathbf{k}}^{\text{ext}}$ are the eigenvalues of the matrix R^{ext} .

In Quasi-Monte Carlo (QMC) convergence theory (see for example [13], [12]) it is important to have good estimates for $\|\mathbf{B}_{\mathbf{k}}\|_{\infty}$. More precisely, arranging the $\Lambda_{\mathbf{k}}^{\text{ext}}$ in non-increasing order, it is important to study the rate of decay of the resulting sequence. In order to obtain some insight into this question we first recall that the $\Lambda_{\mathbf{k}}^{\text{ext}}$ depend on both the regular mesh diameter h_0 and the extension length ℓ or equivalently, on h_0 and m (see (2.1)). Since $m = \ell h_0^{-1}$, the dimension s given by (1.7) then grows if either h_0 decreases or ℓ increases (or both). To indicate the dependence on these two parameters, in this section we write variously

$$\Lambda_{\mathbf{k}}^{\text{ext}} = \Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell) = \Lambda_{s, \mathbf{k}}^{\text{ext}}.$$

In order to get insight into the asymptotic behaviour of $\Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell)$, we study first the spectrum of the continuous periodic covariance integral operator defined by

$$\mathcal{R}^{\text{ext}} v(\mathbf{x}) := \int_{[0, 2\ell]^d} \rho^{\text{ext}}(\mathbf{x} - \mathbf{x}') v(\mathbf{x}') d\mathbf{x}', \quad \mathbf{x} \in [0, 2\ell]^d,$$

where ρ^{ext} is defined in (2.4). This operator is a continuous analogue of the matrix R^{ext} defined in (2.6) and the eigenvalues of each are closely related as we shall discuss below.

The operator \mathcal{R}^{ext} is a compact operator on the space of 2ℓ -periodic continuous functions on \mathbb{R}^d , and so it has a discrete spectrum with the only accumulation point at the origin. Since \mathcal{R}^{ext} is a periodic convolution operator, it is easily verified that its eigenvalues and (normalised) eigenfunctions (which depend on ℓ) are

$$\lambda_{\mathbf{k}}^{\text{ext}}(\ell) = \int_{[0,2\ell]^d} \rho^{\text{ext}}(\mathbf{x}) \exp(-2\pi i \boldsymbol{\xi}_{\mathbf{k}} \cdot \mathbf{x}) d\mathbf{x} = \int_{[-\ell,\ell]^d} \rho(\mathbf{x}) \exp(-2\pi i \boldsymbol{\xi}_{\mathbf{k}} \cdot \mathbf{x}) d\mathbf{x} \quad (3.1)$$

$$\text{and } v_{\mathbf{k}}(\mathbf{x}) = (2\ell)^{-d/2} \exp(2\pi i \boldsymbol{\xi}_{\mathbf{k}} \cdot \mathbf{x}), \quad \text{and where } \boldsymbol{\xi}_{\mathbf{k}} = \mathbf{k}/(2\ell) \quad \text{and } \mathbf{k} \in \mathbb{Z}^d.$$

Here we used the fact that ρ^{ext} is 2ℓ -periodic and that ρ^{ext} and ρ coincide on $[-\ell, \ell]^d$. The eigenvalues $\lambda_{\mathbf{k}}^{\text{ext}}(\ell)$ are real but not necessarily positive, since ρ^{ext} , unlike ρ , may not be positive definite (for the definition see §2.2).

The close relationship between the eigenvalues of the continuous operator and the matrix eigenvalues is shown by the following theorem.

Theorem 3.1 *For fixed $\ell \geq 1$ and fixed $\mathbf{k} \in \overline{\mathbb{Z}}_m^d$, where $m = \ell/h_0$, the matrix eigenvalues $\Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell)$, weighted by h_0^d , converge to $\lambda_{\mathbf{k}}^{\text{ext}}(\ell)$ as $h_0 \rightarrow 0$:*

$$h_0^d \Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell) \rightarrow \lambda_{\mathbf{k}}^{\text{ext}}(\ell), \quad \text{as } h_0 \rightarrow 0.$$

Proof. The formula for $\Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell)$ given by (2.7), when weighted by h_0^d , can be seen as a rectangle rule approximation of the second integral defining $\lambda_{\mathbf{k}}^{\text{ext}}(\ell)$ in (3.1), with grid spacing of h_0 . Since the integrand in that integral is continuous, the convergence claim holds. \square

From now on we restrict attention to the Matérn case (Example 2.7). Our first result shows that $\lambda_{\mathbf{k}}^{\text{ext}}(\ell)$ is exponentially close to the full range Fourier transform $\widehat{\rho}(\boldsymbol{\xi}_{\mathbf{k}})$ (given in (2.10)), and this holds true uniformly in ℓ and \mathbf{k} .

Lemma 3.2 *For the case of the Matérn kernel, we have*

$$|\lambda_{\mathbf{k}}^{\text{ext}}(\ell) - \widehat{\rho}(\boldsymbol{\xi}_{\mathbf{k}})| \lesssim \lambda^d 5^\nu \nu^{d/2-1} \exp\left(-\sqrt{\frac{\nu}{2}} \frac{\ell}{\lambda}\right), \quad \text{for all } \ell \geq 1 \quad \text{and } \mathbf{k} \in \overline{\mathbb{Z}}_m^d. \quad (3.2)$$

Thus there exist positive constants C_3, C_4 (independent of h_0, ℓ, λ and ν), such that, for all $\varepsilon > 0$ and all $\mathbf{k} \in \overline{\mathbb{Z}}_m^d$,

$$|\lambda_{\mathbf{k}}^{\text{ext}}(\ell) - \widehat{\rho}(\boldsymbol{\xi}_{\mathbf{k}})| \leq \varepsilon, \quad \text{when } \ell/\lambda \geq C_3 \nu^{1/2} + C_4 \nu^{-1/2} \log(1/\varepsilon). \quad (3.3)$$

Proof. From (3.1), (2.10) and (2.21), it is easy to see that

$$|\lambda_{\mathbf{k}}^{\text{ext}}(\ell) - \widehat{\rho}(\boldsymbol{\xi}_{\mathbf{k}})| \leq \int_{\|\mathbf{x}\|_2 \geq \ell} |\rho(\mathbf{x})| d\mathbf{x} \sim \int_{\ell}^{\infty} r^{d-1} |\kappa(r/\lambda)| dr = \lambda^d \int_{\ell/\lambda}^{\infty} r^{d-1} \kappa(r) dr,$$

with κ as in (2.21), and again we use Notation 2.8. Then, using (2.25), we obtain (3.2).

Forcing the right-hand side of (3.2) to be bounded above by ε , rearranging and taking logs, we obtain a sufficient condition of the form

$$\sqrt{\frac{\nu}{2}} \frac{\ell}{\lambda} \geq C_* + \log(1/\varepsilon) + d \log \lambda + \nu \log 5 + (d/2 - 1) \log \nu,$$

for some parameter-independent constant C_* . Recalling that we assume $\lambda \leq 1$, and using $\nu \geq 1/2$ and $\log(\nu)/\nu \lesssim 1$, the sufficiency condition in (3.3) follows. \square

With Lemma 3.2 in mind, we now discuss the asymptotic behaviour of $\widehat{\rho}(\boldsymbol{\xi}_k)$, by making use of the analytic formula (2.22). In order to define an appropriate ordering we make the following definition.

Definition 3.3 (Ordering of the integer lattice) *Since \mathbb{Z}^d is countable we can write its elements as a sequence $\{\mathbf{k}(j) : j = 1, 2, \dots\}$, such that $\mathbf{k}(1) = \mathbf{0} \in \mathbb{Z}^d$ and such that the sequence $\{\|\mathbf{k}(j)\|_2 : j = 1, 2, \dots\}$ is non-decreasing. This ordering is not unique.*

Theorem 3.4 *For the case of the Matérn kernel, we have*

$$0 < \widehat{\rho}(\boldsymbol{\xi}_{\mathbf{k}(j)}) \lesssim \lambda^{-2\nu} (\nu \ell^2)^{\nu+d/2} j^{-(1+2\nu/d)}, \quad j = 1, 2, \dots \quad (3.4)$$

Proof. By (2.22) and the definition $\boldsymbol{\xi}_k = \mathbf{k}/(2\ell)$, we have

$$0 < \widehat{\rho}(\boldsymbol{\xi}_k) \sim \lambda^d \nu^\nu \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} (\nu + \pi^2 \lambda^2 \|\mathbf{k}\|_2^2 / (2\ell^2))^{-(\nu+d/2)}.$$

Now by (2.26) we have (for even d), $\Gamma(\nu + d/2) \lesssim \nu^{d/2} \Gamma(\nu)$. Moreover the same estimate also holds for d odd, as can be seen by employing the first equation in (2.27), and then the estimate $\Gamma(\nu + 1/2) \lesssim \nu^{1/2} \Gamma(\nu)$ (from [15, eq. (1.3)]). Hence we have the upper bound

$$\begin{aligned} \widehat{\rho}(\boldsymbol{\xi}_k) &\lesssim \lambda^d \nu^{\nu+d/2} (\nu + \pi^2 \lambda^2 \|\mathbf{k}\|_2^2 / (2\ell^2))^{-(\nu+d/2)} \lesssim \lambda^d (\nu \ell^2)^{\nu+d/2} (\lambda \|\mathbf{k}\|_2)^{-(2\nu+d)} \\ &= \lambda^{-2\nu} (\nu \ell^2)^{\nu+d/2} \|\mathbf{k}\|_2^{-(2\nu+d)}. \end{aligned}$$

Now, to complete the proof of (3.4), we shall show that

$$\|\mathbf{k}(j)\|_2 \sim j^{1/d}. \quad (3.5)$$

To obtain (3.5), for each j , define the set $T(j) := \{\mathbf{k} \in \mathbb{Z}^d : \|\mathbf{k}\|_2 \leq \|\mathbf{k}(j)\|_2\}$. Then $T(j)$ is a superset and a subset of two cubes:

$$[-d^{-1/2} \|\mathbf{k}(j)\|_2, d^{-1/2} \|\mathbf{k}(j)\|_2]^d \subseteq T(j) \subseteq [-\|\mathbf{k}(j)\|_2, \|\mathbf{k}(j)\|_2]^d.$$

(The right inclusion follows from the definition of $T(j)$. On the other hand if \mathbf{k} lies in the left-most cube, then $\|\mathbf{k}\|_\infty \leq d^{-1/2} \|\mathbf{k}(j)\|_2$ which implies $\|\mathbf{k}\|_2 \leq \|\mathbf{k}(j)\|_2$, i.e., $\mathbf{k} \in T(j)$.) Then it follows that $\#T(j) \sim \|\mathbf{k}(j)\|_2^d$. But also by definition $T(j)$ contains $j - 1$ elements of \mathbb{Z}^d . Hence (3.5) follows. \square

Combining Lemma 3.2 and Theorem 3.4 with Theorem 2.9, we obtain the following criterion which simultaneously guarantees the positivity of all the eigenvalues of R^{ext} and provides an upper bound for $\lambda_{\mathbf{k}(j)}(\ell)$, explicit in the parameters h_0, ℓ, λ and ν .

Corollary 3.5 *For the Matérn kernel, suppose*

$$\ell/\lambda \geq \max\{C_1, C_3\nu^{1/2}\} + \max\{C_2\nu^{1/2}, C_4\nu^{-1/2}\} \log(\max\{\lambda/h_0, \nu^{1/2}\}).$$

Then $\Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell) \geq 0$ for all $\mathbf{k} \in \overline{\mathbb{Z}}_m^d$ and also

$$\lambda_{\mathbf{k}(j)}^{\text{ext}}(\ell) \leq \min\{h_0/\lambda, \nu^{-1/2}\} + \lambda^{-2\nu} (\nu\ell^2)^{\nu+d/2} j^{-(1+2\nu/d)}, \quad j = 1, 2, \dots \quad (3.6)$$

Discussion 3.6 *In the paper [12] the authors of this paper analyse the convergence of Quasi-Monte Carlo methods for uncertainty quantification of certain PDEs with random coefficients, realisations of which are computed using the circulant embedding technique described here. A dimension-independent convergence estimate for the QMC method is proved there under the assumption that given d and covariance function ρ , there exist $p < 1$ and $C > 0$ such that for all integers m_0 , with m chosen as in Algorithm 1 and $s = (2m)^d$, we have*

$$\sum_{\mathbf{k} \in \overline{\mathbb{Z}}_m^d} \left(\frac{\Lambda_{s, \mathbf{k}}^{\text{ext}}}{s} \right)^{p/2} \leq C. \quad (3.7)$$

If $\Lambda_{s, j}^{\text{ext}}$ for $j = 1, \dots, s$ denotes the eigenvalues $\Lambda_{s, \mathbf{k}}^{\text{ext}}$ reordered so as to be nonincreasing, then a sufficient condition for this to hold is that for some $\beta > 1$ and $C > 0$, independent of j and s , there holds

$$\sqrt{\frac{\Lambda_{s, j}^{\text{ext}}}{s}} \leq C j^{-\beta} \quad \text{for } j = 1, \dots, s, \quad (3.8)$$

in which case (3.7) holds for p any number in the open interval $(1/\beta, 1)$. Based on the known rate of decay of the analogous Karhunen-Loève eigenvalues (e.g., [18], [13]) and the fact that the second term in (3.6) decays with the same rate (and bearing in mind the result of Theorem 3.1), we conjecture that

$$\textbf{Conjecture:} \quad \text{For the Matérn kernel, (3.8) holds with } \beta = \frac{1 + 2\nu/d}{2}. \quad (3.9)$$

Numerical experiments in the next section support this conjecture. Thus we need to assume that $\nu > d/2$ in order to ensure that $\beta > 1$ and $p < 1$. (If $\nu \leq d/2$ then there is no predicted convergence rate for the QMC algorithm. A similar theoretical barrier was detected in [13]).

The conjecture that $\sqrt{\Lambda_{s, j}^{\text{ext}}/s} \leq C j^{-(1+2\nu/d)/2}$ remains a conjecture in spite of the results in Lemma 3.2 and Theorem 3.4. That is because these results concern the continuous eigenvalues $\lambda_{\mathbf{k}}^{\text{ext}}(\ell)$ and the result which connects these with the eigenvalues $\Lambda_{\mathbf{k}}^{\text{ext}}(h_0, \ell)$ (namely Theorem 3.1) does not provide enough information about the

small eigenvalues. Indeed it is well known, and demonstrated in the numerical experiments in the next section, that the small matrix eigenvalues depart significantly from the corresponding eigenvalues $\lambda_{\mathbf{k}}^{\text{ext}}(\ell)$ of the integral operator.

A sufficient condition for (3.8), and hence for (3.7), is that (3.8) holds not for all eigenvalues but for some fixed fraction of the eigenvalues, say α , starting from the largest. Suppose, for example, that for some fixed $\alpha \in (0, 1)$ we have

$$\sqrt{\frac{\Lambda_{s,j}^{\text{ext}}}{s}} \leq Cj^{-\beta} \quad \text{for } j = 1, \dots, \lceil \alpha s \rceil.$$

Then because the eigenvalues are ordered, we also have, for $j = \lceil \alpha s \rceil + 1, \dots, s$,

$$\sqrt{\frac{\Lambda_{s,j}^{\text{ext}}}{s}} \leq C\lceil \alpha s \rceil^{-\beta} = Cj^{-\beta} \left(\frac{j}{\lceil \alpha s \rceil} \right)^{\beta} \leq \frac{C}{\alpha^{\beta}} j^{-\beta},$$

so that the sufficiency condition (3.8) is satisfied for all $j = 1, \dots, s$ with an appropriately larger value of the constant.

4 Numerical Experiments

In this section we perform numerical experiments illustrating the theoretical results given above, for the Matérn covariance in 2D and 3D. In all experiments we set the variance $\sigma^2 = 1$. In §4.1, we illustrate the positive definiteness results of Theorem 2.9. Then, in §4.2, we illustrate the actual decay of the eigenvalues $\Lambda_{s,j}^{\text{ext}}$ and compare it to (3.9).

4.1 Positive definiteness

We investigate experimentally the minimal value of ℓ needed to ensure that the extended circulant matrix R^{ext} is positive definite. Recall that Theorem 2.3 guarantees the existence of such an ℓ and Theorems 2.9 and 2.11 quantify its behaviour in the Matérn case.

The plots in Figure 1 show the behaviour of ℓ as a function of $\log h_0^{-1} = \log m_0$ for various choices of d , $\nu < \infty$ and λ . They clearly show that ℓ depends linearly on $\log m_0$, for m_0 large enough. They also indicate that ℓ gets smaller if either λ or ν gets smaller, all in accordance with Theorem 2.9. For fixed d , h_0 and λ , Theorem 2.9 gives a lower bound on ℓ which grows like $\nu^{1/2}$. We observe this behaviour clearly for $d = 2$, but the growth with respect to ν is a bit slower for $d = 3$. (See Figure 2, which plots $\log \ell$ against $\log \nu$.) The small triangle embedded in each graph indicates $\mathcal{O}(\nu^{1/2})$ growth.

Table 1 illustrates the result of Theorem 2.11 which covers the case $\nu = \infty$. Here we tabulate the value of ℓ needed to ensure positive definiteness of R^{ext} , when decreasing λ and keeping $m_0\lambda$ fixed at 8. Theorem 2.11 gives a bound which decreases linearly in λ , until the minimum $\ell = 1$ is reached. This behaviour is exactly as

λ	m_0	ℓ ($d = 2$)	ℓ ($d = 3$)
1	8	8	9
0.5	16	4	4.5
0.25	32	2	2.25
0.125	64	1	1.125

Table 1: Illustration of Theorem 2.11: Values of ℓ required to ensure positive definiteness of R^{ext} in the case $\nu = \infty$, with λ decreasing and $m_0\lambda = 8$ fixed.

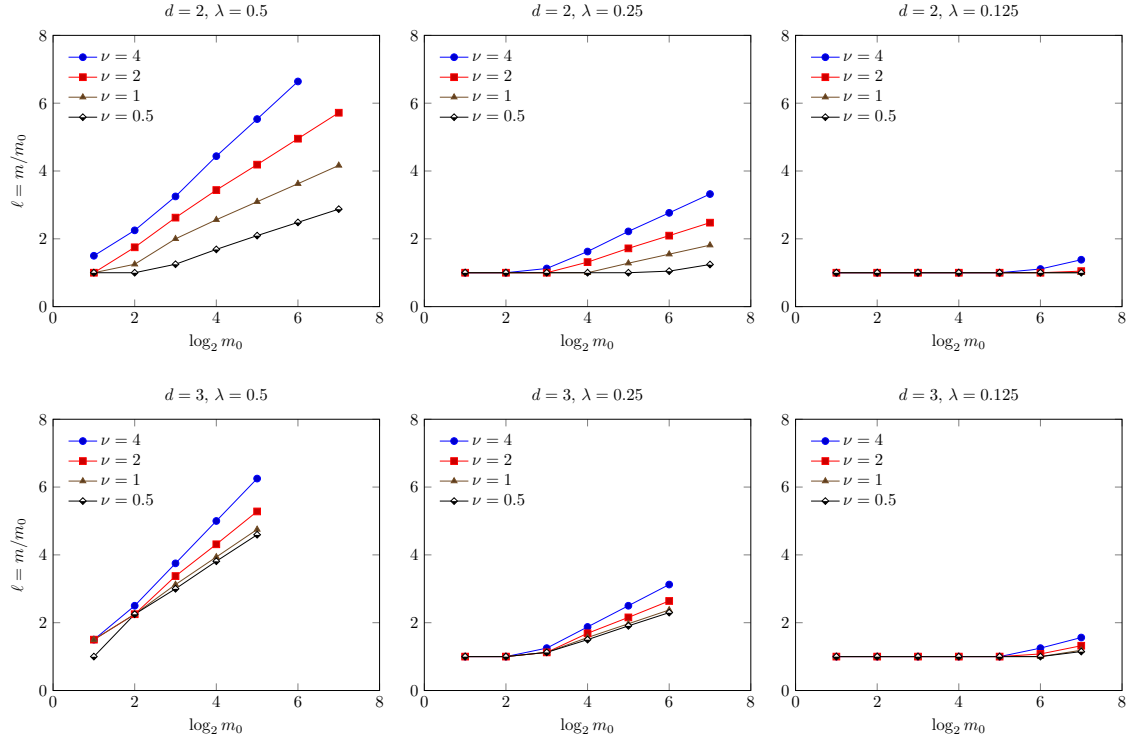


Figure 1: Illustration of Theorem 2.9: Graphs of the minimum needed value of ℓ to obtain positive definiteness against $\log_2 h_0^{-1} = \log_2 m_0$ for different choices of d , λ and ν . The graphs show a linear relationship with lower values of ℓ for smaller λ and for smaller ν .

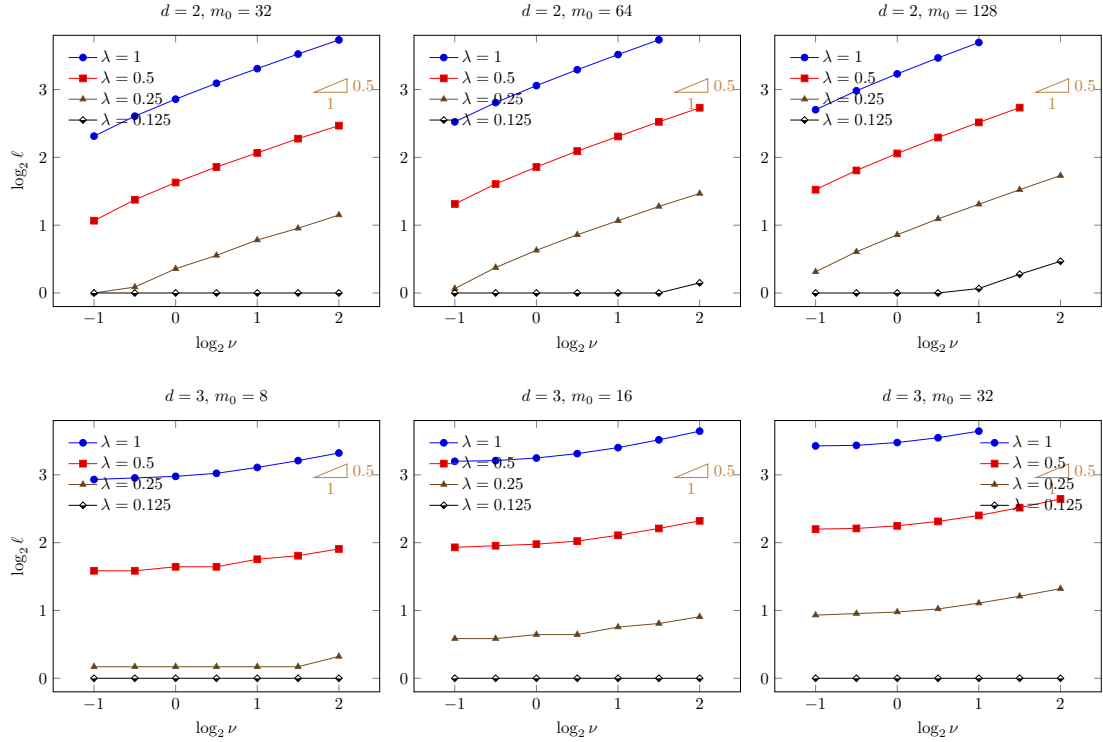


Figure 2: Illustration of Theorem 2.9: Log-log graphs of the minimum needed value of ℓ to obtain positive definiteness against ν for various choices of d , m_0 and λ . The small triangle depicts a gradient of 0.5.

observed in Table 1. In this case R^{ext} has many very small eigenvalues and we deem it to be positive definite when all eigenvalues are positive, ignoring those eigenvalues which are less than 10^{-13} in modulus.

4.2 Eigenvalue decay

Here we perform experiments to verify the decay conjecture (3.9). In Figures 3 and 4 we present log-log plots of $\sqrt{\Lambda_{s,j}^{\text{ext}}/s}$ against the index j for various choices of d , ν and λ . For decreasing h_0 (i.e., increasing m_0) we determine the minimum ℓ to achieve positive definiteness and make the plot for each case. If the conjecture (3.9) holds, then we expect to see a polynomial decay of rate $-(1 + 2\nu/d)/2$, as shown in the slope triangle in each case. (The convergence may tail off eventually; note that Theorem 3.1 does not provide an explicit convergence rate.) We see that the computations closely follow the prediction of (3.9). Moreover, to have the constant C in (3.7) for the minimal embedding to be absolutely bounded independent of h_0 (and thus also independent of s , m and ℓ) we observe the lines to get closer together for increasing m_0 .

Thus, while the numerical evidence supports conjecture (3.9), it remains an interesting open problem to prove it.

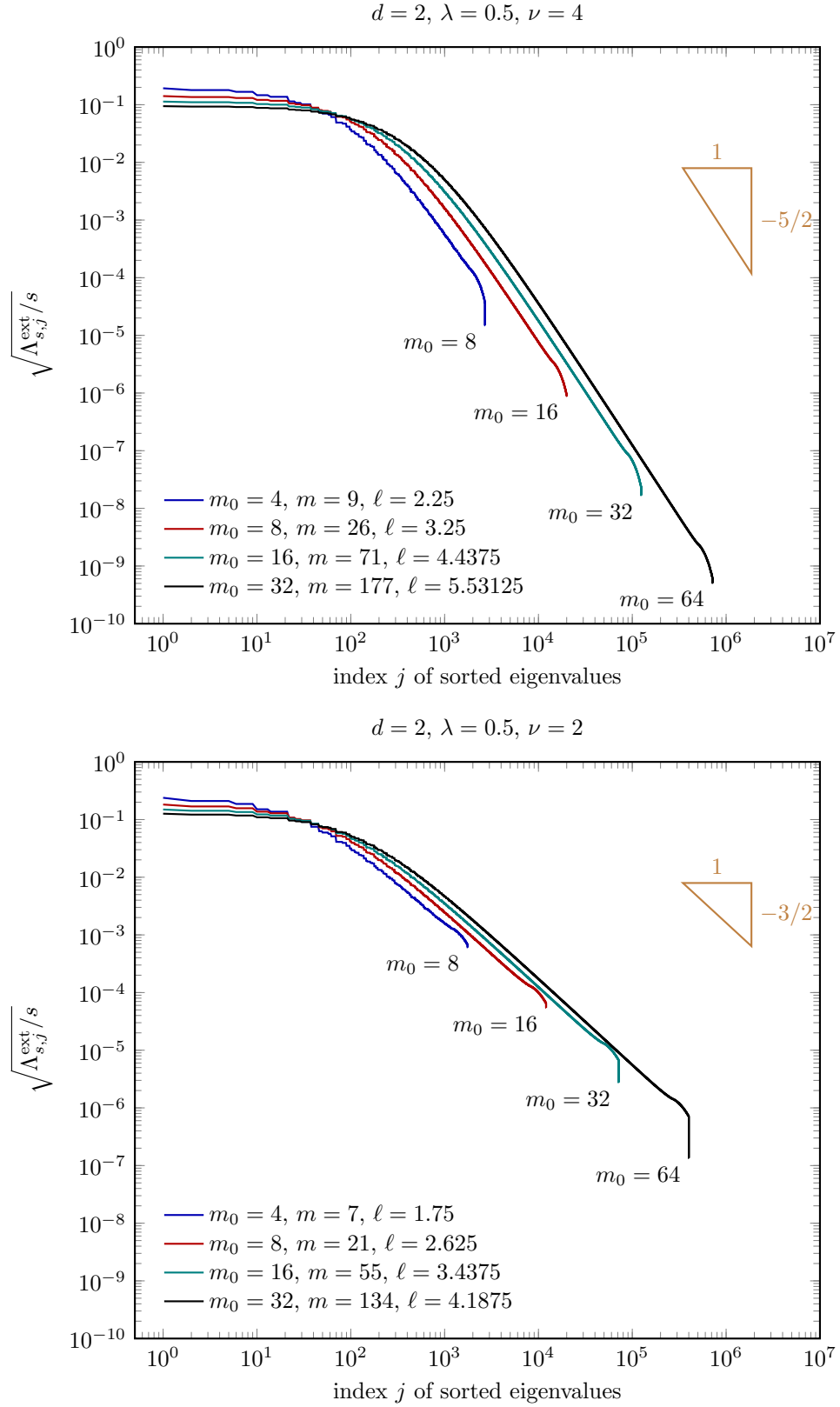


Figure 3: Loglog plot of the decay of the eigenvalues in case of the minimal embedding for $d = 2$, $\lambda = 0.5$, and $\nu = 4$ (top) and $\nu = 2$ (bottom). The expected decay rates (3.9) are marked by the slope triangles.

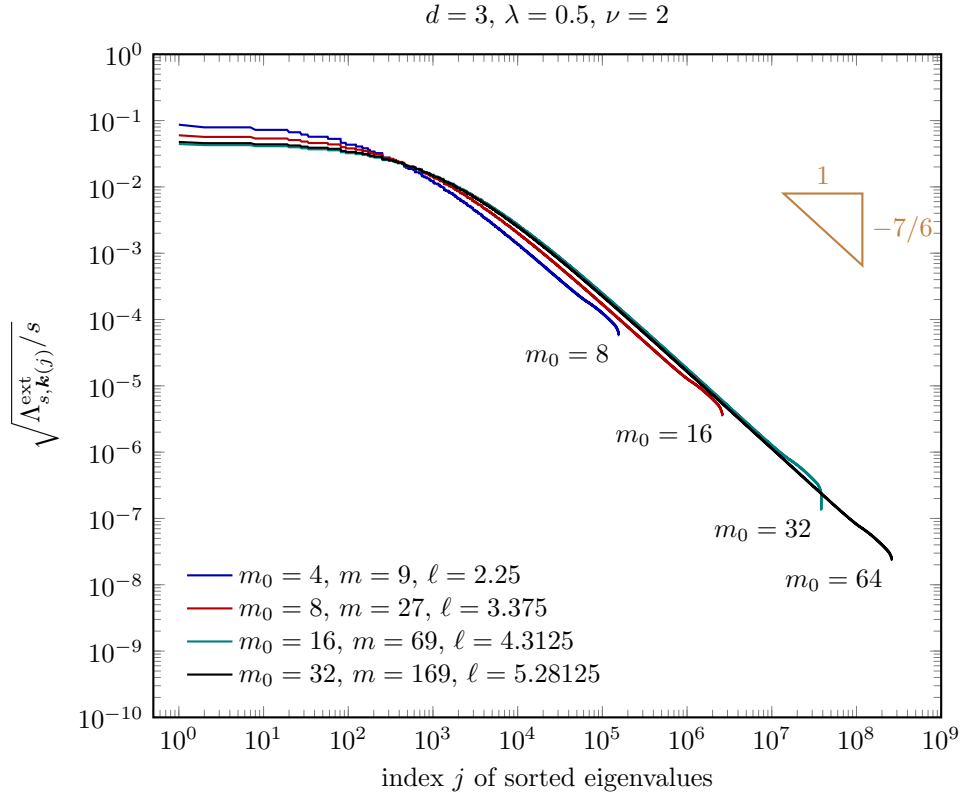


Figure 4: Loglog plot of the decay of the eigenvalues in case of the minimal embedding for $d = 3$, $\lambda = 0.5$ and $\nu = 2$. The expected decay rate (3.9) is marked by the slope triangle.

A Sampling theorem

The following result is a d -dimensional version of what is commonly called a sampling theorem. (See [4] for a 1-dimensional version.)

Theorem A.1 *Let $\rho \in L^1(\mathbb{R}^d)$ be real-valued and symmetric, with $\hat{\rho} \in L^1(\mathbb{R}^d)$. Suppose also that for some $h > 0$,*

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} |\rho(h\mathbf{k})| < \infty. \quad (\text{A.1})$$

Then

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} \rho(h\mathbf{k}) \exp(-2\pi i h \mathbf{k} \cdot \boldsymbol{\xi}) = \frac{1}{h^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \hat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right) \quad \text{for almost all } \boldsymbol{\xi} \in \mathbb{R}^d. \quad (\text{A.2})$$

If $\hat{\rho}$ is everywhere positive, then for all $\boldsymbol{\xi} \in \mathbb{R}^d$

$$\begin{aligned} \sum_{\mathbf{k} \in \mathbb{Z}^d} \rho(h\mathbf{k}) \exp(-2\pi i h \mathbf{k} \cdot \boldsymbol{\xi}) &\geq \operatorname{ess\,inf}_{\boldsymbol{\zeta} \in [-\frac{1}{2h}, \frac{1}{2h}]^d} \frac{1}{h^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \hat{\rho}\left(\boldsymbol{\zeta} + \frac{\mathbf{r}}{h}\right) \\ &\geq \max_{\mathbf{r} \in \mathbb{Z}^d} \min_{\boldsymbol{\zeta} \in [-\frac{1}{2}, \frac{1}{2}]^d} \frac{1}{h^d} \hat{\rho}\left(\frac{\boldsymbol{\zeta} + \mathbf{r}}{h}\right) > 0. \end{aligned} \quad (\text{A.3})$$

Proof. Note first that ρ and $\hat{\rho}$ are both continuous because of the assumed integrability of $\hat{\rho}$ and ρ respectively. Moreover $\hat{\rho}$ is real because of the assumed symmetry of ρ . The infinite sum on the left-hand side of (A.2) is a well-defined continuous function of $\boldsymbol{\xi} \in \mathbb{R}^d$, because the convergence is absolute and uniform by virtue of (A.1).

Now consider the right-hand side of (A.2), writing it for convenience as

$$g_h(\boldsymbol{\xi}) := \frac{1}{h^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \hat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right), \quad \boldsymbol{\xi} \in \mathbb{R}^d.$$

We will show that the function so defined is locally integrable. It is also manifestly $1/h$ periodic in each coordinate direction, since $g_h(\boldsymbol{\xi} + \mathbf{q}/h) = g_h(\boldsymbol{\xi})$ for all $\mathbf{q} \in \mathbb{Z}^d$.

To show the local integrability of g_h we will show that the sum defining g_h converges absolutely to an integrable function on $[0, 1/h]^d$. Letting Ξ denote any finite subset of \mathbb{Z}^d , we have

$$\begin{aligned} \int_{[0, 1/h]^d} \left| \sum_{\mathbf{r} \in \Xi} \hat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right) \right| d\boldsymbol{\xi} &\leq \int_{[0, 1/h]^d} \sum_{\mathbf{r} \in \Xi} \left| \hat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right) \right| d\boldsymbol{\xi} = \sum_{\mathbf{r} \in \Xi} \int_{[0, 1/h]^d} \left| \hat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right) \right| d\boldsymbol{\xi} \\ &= \sum_{\mathbf{r} \in \Xi} \int_{r_1/h}^{(1+r_1)/h} \cdots \int_{r_d/h}^{(1+r_d)/h} |\hat{\rho}(\boldsymbol{\xi})| d\boldsymbol{\xi} \leq \int_{\mathbb{R}^d} |\hat{\rho}(\boldsymbol{\xi})| d\boldsymbol{\xi} < \infty, \end{aligned}$$

where the last step follows from the assumed integrability of $\hat{\rho}$. It then follows from the dominated convergence theorem that g_h is integrable on $[0, 1/h]^d$.

We now show that the left-hand side of (A.2) is just the Fourier series of the integrable $(1/h)$ -periodic function g_h . The Fourier coefficients of g_h are given by

$$\widehat{g}_h(\mathbf{k}) = h^d \int_{-1/(2h)}^{1/(2h)} \cdots \int_{-1/(2h)}^{1/(2h)} \left(\frac{1}{h^d} \sum_{\mathbf{r} \in \mathbb{Z}^d} \widehat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right) \right) \exp(2\pi i h \mathbf{k} \cdot \boldsymbol{\xi}) d\boldsymbol{\xi},$$

where the unconventional sign in the exponent is valid because g_h is real and symmetric. The order of integration and summation can be interchanged by a second application of the dominated convergence argument, to give

$$\begin{aligned} \widehat{g}_h(\mathbf{k}) &= \sum_{\mathbf{r} \in \mathbb{Z}^d} \int_{-1/(2h)}^{1/(2h)} \cdots \int_{-1/(2h)}^{1/(2h)} \widehat{\rho}\left(\boldsymbol{\xi} + \frac{\mathbf{r}}{h}\right) \exp(2\pi i h \mathbf{k} \cdot \boldsymbol{\xi}) d\boldsymbol{\xi} \\ &= \sum_{\mathbf{r} \in \mathbb{Z}^d} \int_{(-1/2+r_1)/h}^{(1/2+r_1)/h} \cdots \int_{(-1/2+r_d)/h}^{(1/2+r_d)/h} \widehat{\rho}(\boldsymbol{\xi}) \exp(2\pi i h \mathbf{k} \cdot \boldsymbol{\xi}) d\boldsymbol{\xi} \\ &= \int_{\mathbb{R}^d} \widehat{\rho}(\boldsymbol{\xi}) \exp(2\pi i h \mathbf{k} \cdot \boldsymbol{\xi}) d\boldsymbol{\xi} = \rho(h\mathbf{k}), \end{aligned}$$

where in the last step we used the fact that the inverse Fourier transform of $\widehat{\rho}$ recovers ρ because of the integrability of ρ . Thus the left-hand side of (A.2) is just the Fourier series of g_h , as asserted. From this it follows that the integrable function g_h is equal almost everywhere to the continuous function on the left-hand side, so completing the proof of (A.2).

If $\widehat{\rho}$ is a positive function it follows that the left-hand side of (A.2) is bounded below by the essential infimum of $g_h(\boldsymbol{\xi})$ over $\boldsymbol{\xi}$, which in turn can be bounded below by retaining only the largest term in the sum, which is positive. \square

References

- [1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*, Dover, New York, 1965.
- [2] R.J. Adler, *The Geometry of Random Fields*, Wiley, London, 1981.
- [3] M. Bachmayr, A. Cohen and G. Migliorati, Representations of Gaussian random fields and approximation of elliptic PDEs with lognormal coefficients, *J. Fourier Anal. Appl.*, published online 29 March 2017.
- [4] D.C. Champeney, *A Handbook of Fourier Transforms*, Cambridge University Press, Cambridge, 1987.
- [5] G. Chan and A.T.A. Wood, Simulation of stationary Gaussian processes in $[0, 1]^d$, *J. Comput. Graph. Stat.*, **3**, 409–432, 1994.
- [6] G. Chan and A.T.A. Wood, Algorithm AS 312: An Algorithm for simulating stationary Gaussian random fields, *Appl. Stat. – J. Roy. St. C*, **46**, 171–181, 1997.

- [7] C.R. Dietrich and G.H. Newsam, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix, *SIAM J. Sci. Comput.*, **18**, 1088–1107, 1997.
- [8] M. Eiermann, O.G. Ernst and E. Ullmann, Computational aspects of the stochastic finite element method, *Comput. Visual. Sci.*, **10**(1), 3–15, 2007.
- [9] M. Feischl, F.Y. Kuo and I.H. Sloan, Fast random field generation with H -matrices, to appear in *Numer. Math.*, 2018.
- [10] R.G. Ghanem and P.D. Spanos, *Stochastic Finite Elements*, Dover, 1991.
- [11] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl and I.H. Sloan, Quasi-Monte carlo methods for elliptic PDEs with random coefficients and applications, *J. Comput. Phys.*, **230**, 3668–3694, 2011.
- [12] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl and I.H. Sloan, Circulant embedding with QMC – analysis for elliptic PDEs with lognormal coefficients, *Preprint arXiv:1710.09254*, 25 October 2017, available at <https://arxiv.org/abs/1710.09254>
- [13] I.G. Graham, F.Y. Kuo, J.A. Nicholls, R. Scheichl, C. Schwab and I.H. Sloan, Quasi-Monte Carlo finite element methods for elliptic PDEs with log-normal random coefficients, *Numer. Math.*, **131**, 329–368, 2015.
- [14] H. Harbrecht, M. Peters and M. Siebenmorgen, Efficient approximation of random fields for numerical application, *Numer. Linear Algebr.*, **22**, 596–617, 2015.
- [15] D. Kershaw, Some extensions of W. Gautschi’s inequalities for the gamma function, *Math. Comput.*, **41**, 607–611, 1983.
- [16] B. N. Khoromskij, A. Litvinenko and H. G. Matthies, Application of hierarchical matrices for computing the Karhunen-Loève expansion, *Computing*, **84**(1-2), 49–67, 2009.
- [17] D.P. Kroese and Z.I. Botev, Spatial Process Generation. In: V. Schmidt (Ed.). *Lectures on Stochastic Geometry, Spatial Statistics and Random Fields*, Volume II: Analysis, Modeling and Simulation of Complex Structures, Springer-Verlag, Berlin, 2014.
- [18] G. Lord, C. Powell, T. Shardlow, *An Introduction to Computational Stochastic PDEs*, Cambridge University Press, Cambridge, 2014.
- [19] R.L. Naff, D.F. Haley, and E.A. Sudicky, High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media 1. Methodology and flow results, *Water Resour. Res.*, **34**, 663–677, 1998.
- [20] R.L. Naff, D.F. Haley, and E.A. Sudicky, High-resolution Monte Carlo simulation of flow and conservative transport in heterogeneous porous media 2. Transport Results, *Water Resour. Res.*, **34**, 679–697, 1998.

- [21] C. Schwab and R.A. Todor, Karhunen-Loève approximation of random fields by generalized fast multipole methods, *J. Comput. Phys.*, **217** 100–122, (2006)
- [22] H. Wendland, *Scattered Data Approximation*, Cambridge University Press, Cambridge, 2005.